

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/498,556

DATE: 06/12/2001

TIME: 15:06:48

Input Set : N:\Crif3\RULE60\09498556.txt

Output Set: N:\CRF3\06122001\I498556.raw

ENTERED

SEQUENCE LISTING

5 (1) GENERAL INFORMATION:

8 (i) APPLICANT: Vlasuk, George Phillip
 9 Stanssens, Patrick Eric Hugo
 10 Messens, Boris Hilda Lieven
 11 Lauwerijs, Marc Joset
 12 Laroche, Yves Rene
 13 Despers, Laurent Stephane
 14 Gensemans, Yannick Georges Jozef
 15 Hoyle, Matthew
 16 Berquin, Peter W.

17 (ii) TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
 18 INHIBITORS AND ANTICOAGULANT
 19 PROTEIN

20 (iii) NUMBER OF SEQUENCES: 156

21 (iv) CORRESPONDENCE ADDRESS:

22 (A) ADDRESSEE: Lyon & Lyon
 23 (B) STREET: 618 West Fifth Street
 24 Suite 4700
 25 (C) CITY: Los Angeles
 26 (D) STATE: California
 27 (E) COUNTRY: U.S.A.
 28 (F) ZIP: 90011

29 (v) COMPUTER READABLE FORM:

30 (A) MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 31 storage
 32 (B) COMPUTER: IBM Compatible
 33 (C) OPERATING SYSTEM: IBM P.C. DOS 5.0
 34 (D) SOFTWARE: Word Perfect 5.1

35 (vi) CURRENT APPLICATION DATA:

C--> 36 (A) APPLICATION NUMBER: US/09/498,556
 C--> 37 (B) FILING DATE: 04-Feb-2000

38 (vii) PRIOR APPLICATION DATA:

39 (A) APPLICATION NUMBER: 08-800,455
 40 (B) FILING DATE: April 17, 1997
 41 (A) APPLICATION NUMBER: PCT/US95/13231
 42 (B) FILING DATE: October 1, 1995
 43 (A) APPLICATION NUMBER: 08-456,399
 44 (B) FILING DATE: June 5, 1995
 45 (A) APPLICATION NUMBER: 08-456,397
 46 (B) FILING DATE: June 5, 1995
 47 (A) APPLICATION NUMBER: 08-460,380
 48 (B) FILING DATE: June 5, 1995
 49 (A) APPLICATION NUMBER: 08-461,965
 50 (B) FILING DATE: June 5, 1995
 51 (A) APPLICATION NUMBER: 06/326,110
 52 (B) FILING DATE: October 18, 1994

RAW SEQUENCE LISTING

DATE: 06/12/2001

PATENT APPLICATION: US/09/498,556

TIME: 15:06:48

Input Set : N:\Crif3\RULE60\09498556.txt

Output Set : N:\CRF3\06122001\I498556.raw

```

80 (viii) ATTORNEY/AGENT INFORMATION:
81 (A) NAME: ZIGGS, SUZANNE L.
82 (B) REGISTRATION NUMBER: 30,158
83 (C) REFERENCE DOCKET NUMBER: 216/270
84 (ix) TELECOMMUNICATION INFORMATION:
85 (A) TELEPHONE: (213) 489-1600
86 (B) TELEFAX: (213) 955-0440
87 (C) TELEX: 67-5110
88 (2) INFORMATION FOR SEQ ID NO: 1:
89 (i) SEQUENCE CHARACTERISTICS:
90 (A) LENGTH: 141 base pairs
91 (B) TYPE: nucleic acid
92 (C) STRANDEDNESS: single
93 (D) TOPOLOGY: linear
94 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
95 AAGGCATACCT CAGATCTGGC TGAAGATGAA TGGCTCGACG ACTGTGGAAC TCAGAAGCCA 60
96 TGGGAGGCCA AATCAATGA GGAATTCCTT GAGGAGGAAG ATCCGATATG CCGCTCACGT 120
97 GGTGTTTAT TGCCTCTGC TTGGCTATTC AAAGACGGAT TCTACAGAGA CACGGTGATC 180
98 GGGCACTGTG TTAAGGAAA AGGATGGAAC CAACATCAGA TTATACATGT CTGA 234
99 (2) INFORMATION FOR SEQ ID NO: 2:
100 (i) SEQUENCE CHARACTERISTICS:
101 (A) LENGTH: 128 base pairs
102 (B) TYPE: nucleic acid
103 (C) STRANDEDNESS: single
104 (D) TOPOLOGY: linear
105 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
106 AAGGCATACCT CAGATCTGGC TGAAGATGAA TGGCTCGACG TGTGTGGAAC TAAGAAGCCA 60
107 TGGGAGGCCA AATCAATGA GGAATTCCTT GAGGAGGAAG ATCCGATATG CCGCTCACGT 120
108 CCGGGTCCCG CAGCTTGGT ATGCAAGGAC GGATTCTACA GAGACACGGT GATCGGCGAC 180
109 TGTGTAAAGG AAGAAAGATG CAGCAGATAT GAGATTATAC ATGTCTGA 228
110 (2) INFORMATION FOR SEQ ID NO: 3:
111 (i) SEQUENCE CHARACTERISTICS:
112 (A) LENGTH: 161 base pairs
113 (B) TYPE: nucleic acid
114 (C) STRANDEDNESS: single
115 (D) TOPOLOGY: linear
116 (vi) ORIGINAL SOURCE:
117 (A) ORGANISM: Anrylptoma caninum
118 (ix) FEATURE:
119 (A) NAME/KEY: Coding Sequence
120 (B) LOCATION: 25...301
121 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
122 GAATTCGGCT AACTACTAAC A ATG AAG ATG CTT TAC GCT ATC GCT ATA ATG 51
123 Met Lys Met Leu Tyr Ala Ile Ala Ile Met
124 1 5 10
125 TTT CTC CTC GTA TTA TTA TGA AAG GCA AGA ACA GTG AGG AAG GCA TAC 99
126 Phe Leu Leu Val Ser Leu Cys Ser Ala Arg Thr Val Arg Lys Ala Tyr
127 15 20 25
128 CCG GAG TGT GGT GAG AAT GAA TAG CTC GAC GAC TGT GGA ACT CAG AAG 147

```

RAW SEQUENCE LISTING

DATE: 06/12/2001

PATENT APPLICATION: US/09/498,556

TIME: 15:06:48

Input Set : N:\Crif3\RULE60\09498556.txt

Output Set: N:\CRF3\06122001\I498556.raw

```

164 Pro Glu Cys Gly Glu Asn Glu Trp Leu Asp Asp Cys Gly Thr Gln Lys
165          30          35          40
167 CCA TGC GAG GCT AAG TGC AAT GAG GAA CCT CCT GAG GAG GAA GAT CCG      195
168 Pro Cys Glu Ala Lys Cys Asn Glu Glu Pro Pro Glu Glu Asp Pro
169          45          50          55
171 ATA TGC GGC TCA GCT GGT TGT TTA TTA CCT CCT GCT TGC GTA TGC AAA      243
172 Ile Cys Arg Ser Arg Gly Cys Leu Leu Pro Pro Ala Cys Val Cys Lys
173          60          65          70
175 GAC GCA TTC TAT AGA GAC AGG GTG ATC GGC GAC TGT GTT AGG GAA GAA      291
176 Asp Gly Phe Tyr Arg Asp Thr Val Ile Gly Asp Cys Val Arg Glu Glu
177          75          80          85          90
179 GAA TGC GAC GAA CAT GAG ATT ATA CAT GTC T GAACGAGAAA GCAACAATAA CC      344
180 Glu Cys Asp Glu His Glu Ile Ile His Val
181          95          100
183 AAAGGTTCCA ACTCTGCTTC TGCAGAAATCG CTAGTIGGAT GTCTCTTTTG CGTCCGAATA      404
185 GTTTTAGTTC AGTTAAGTA AGAAGTCTTG CTGAGAGAGAA TAAAGCTTTC CAACTCC      461
187 (2) INFORMATION FOR SEQ ID NO: 4:
188     (i) SEQUENCE CHARACTERISTICS:
189         (A) LENGTH: 77 amino acids
190         (B) TYPE: amino acid
191         (D) TOPOLOGY: linear
192     (ii) MOLECULE TYPE: peptide
193     (vi) ORIGINAL SOURCE:
194         (A) ORGANISM: Ancylostoma caninum
195     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
196 Lys Ala Tyr Pro Glu Cys Gly Glu Asn Glu Trp Leu Asp Asp
197     1          5          10
198 Cys Gly Thr Gln Lys Phe Cys Glu Ala Lys Cys Asn Glu Glu
199     15          20          25
201 Pro Pro Glu Glu Glu Asp Pro Ile Cys Arg Ser Arg Gly Cys
202     30          35          40
204 Leu Leu Pro Pro Ala Cys Val Cys Lys Asp Gly Phe Tyr Arg
205     45          50          55
207 Asp Thr Val Ile Gly Asp Cys Val Arg Glu Glu Glu Cys Asp
208     60          65          70
209 Gln His Glu Ile Ile His Val
210          75
212 (2) INFORMATION FOR SEQ ID NO: 5:
213     (i) SEQUENCE CHARACTERISTICS:
214         (A) LENGTH: 455 base pairs
215         (B) TYPE: nucleic acid
216         (C) ORGANISMES: single
217         (D) TOPOLOGY: linear
218     (vi) ORIGINAL SOURCE:
219         (A) ORGANISM: Ancylostoma caninum
220     (ix) FEATURE:
221         (A) NAME KEY: Coding Sequence
222         (B) LOCATION: 22...315
223     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/498,556

DATE: 06/12/2001

TIME: 11:06:48

Input Set : N:\Crif3\RULE60\09498556.txt

Output Set: N:\CRF3\06122001\I498556.raw

```

244 GAATTCGCT ACTACTCAAC A ATG AAG ATG CTT TAC GCT ATC GCT ATA AAG 51
245 Met Lys Met Leu Tyr Ala Ile Ala Ile Met
246 1 5 10
248 TTT CTC CTG GCG TCA TTA TGC AGC ACA AGA ACA GTG AGG AAG GCA TAC 99
249 Phe Leu Leu Val Ser Leu Cys Ser Thr Arg Thr Val Arg Lys Ala Tyr
250 15 20 25
252 CCG GAG TGT GGT GAG AAT GAA TGG CTC GTC TGT GGA ACT AAG AAG 147
253 Pro Glu Cys Gly Glu Asn Glu Trp Leu Asp Val Cys Gly Thr Lys Lys
254 30 35 40
256 CCA TGC GAG GTC AAG TGC AGT GAG GAA GAG GAG GAA GAT CCG ATA TAC 195
257 Pro Cys Glu Ala Lys Cys Ser Glu Glu Glu Glu Glu Asp Pro Ile Cys
258 45 50 55
262 CGA TCA TTT TCT TGT CCG GGT CCG GGT GGT TGT TGC GTA TGC GAA GAC GGA 243
263 Arg Ser Phe Ser Cys Pro Gly Pro Ala Ala Cys Val Cys Glu Asp Gly
264 60 65 70
266 TTC TAC AGA GAC ACG GTG ATC GGC GAC TGT GTT AAG GAA GAA GAA TAC 291
267 Phe Tyr Arg Asp Thr Val Ile Gly Asp Cys Val Lys Glu Glu Glu Cys
268 75 80 85 90
270 GAC CAA CAT GAG ATT ATT CAT GTC TGAACGAGAG AGCAGTAATA ACCAAAGGTT C 346
271 Asp Gln His Glu Ile Ile His Val
272 95
274 CAACTTTCGC TTACAAAAT CGCTAGTTGG ATTCTCCTT TCGGTGCGAA TAGTTTTAGT 406
275 TGATATTAAAG TAAACCTCC TGTGAAGAG AATAAAGCTT TCCAACTTC 455
279 (2) INFORMATION FOR SEQ ID NO: 6:
281 (i) SEQUENCE CHARACTERISTICS:
283 (A) LENGTH: 75 amino acids
284 (B) TYPE: amino acid
285 (C) TOPOLOGY: linear
287 (ii) MOLECULE TYPE: peptide
289 (vi) ORIGINAL SOURCE:
291 (A) ORGANISM: Ancylostoma caninum
293 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
295 Lys Ala Tyr Pro Glu Cys Gly Glu Asn Glu Trp Leu Asp Val Cys Gly
296 1 10 15
298 Thr Lys Lys Pro Cys Glu Ala Lys Cys Ser Glu Glu Glu Glu Glu Asp
299 20 25 30
301 Pro Ile Cys Arg Ser Phe Ser Cys Pro Gly Pro Ala Ala Cys Val Cys
302 35 40 45
304 Glu Asp Gly Phe Tyr Arg Asp Thr Val Ile Gly Asp Cys Val Lys Glu
305 50 55 60
307 Glu Glu Cys Arg Gln His Glu Ile Ile His Val
308 65 70 75
310 (2) INFORMATION FOR SEQ ID NO: 7:
312 (i) SEQUENCE CHARACTERISTICS:
314 (A) LENGTH: 81 amino acids
315 (B) TYPE: amino acid
316 (C) TOPOLOGY: linear
318 (ii) MOLECULE TYPE: peptide
320 (vi) ORIGINAL SOURCE:

```

RAW SEQUENCE LISTING

DATE: 06/12/2001

PATENT APPLICATION: US/09/498,556

TIME: 15:06:48

Input Set : N:\Crif3\RULE60\09498556.txt

Output Set: N:\CRF3\06122001\I498556.raw

333 (A) ORGANISM: Ancylostoma caninum
 334 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
 335 Arg Thr Val Ala Lys Ala Tyr Pro Glu Cys Gly Glu Asn Glu Trp Leu
 336 1 10 15
 337 Asp Asp Cys Gly Thr Gln Lys Pro Cys Glu Ala Lys Cys Asn Glu Glu
 338 20 25 30
 339 Pro Pro Glu Glu Glu Asp Pro Ile Cys Arg Ser Arg Gly Cys Leu Leu
 340 35 40 45
 341 Pro Pro Ala Cys Val Cys Lys Asp Gly Phe Tyr Arg Asp Thr Val Ile
 342 50 55 60
 343 Gly Asp Cys Val Arg Glu Glu Glu Cys Asp Gln His Glu Ile Ile His
 344 65 70 75 80
 345 Val

346 (2) INFORMATION FOR SEQ ID NO: 7:

347 (i) SEQUENCE CHARACTERISTICS:

348 (A) LENGTH: 79 amino acids

349 (B) TYPE: amino acid

350 (C) TOPOLOGY: linear

351 (ii) MOLECULE TYPE: peptide

352 (vi) ORIGINAL SOURCE:

353 (A) ORGANISM: Ancylostoma caninum

354 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

355 Arg Thr Val Ala Lys Ala Tyr Pro Glu Cys Gly Glu Asn Glu Trp Leu
 356 1 10 15
 357 Asp Val Cys Gly Thr Lys Lys Pro Cys Glu Ala Lys Cys Ser Glu Glu
 358 20 25 30
 359 Glu Glu Glu Asp Pro Ile Cys Arg Ser Phe Ser Cys Pro Gly Pro Ala
 360 35 40 45
 361 Ala Cys Val Cys Glu Asp Gly Phe Tyr Arg Asp Thr Val Ile Gly Asp
 362 50 55 60
 363 Cys Val Lys Glu Glu Glu Cys Asp Gln His Glu Ile Ile His Val
 364 65 70 75

365 (2) INFORMATION FOR SEQ ID NO: 9:

366 (i) SEQUENCE CHARACTERISTICS:

367 (A) LENGTH: 111 base pairs

368 (B) TYPE: nucleic acid

369 (C) STRANDEDNESS: single

370 (D) TOPOLOGY: linear

371 (vi) ORIGINAL SOURCE:

372 (A) ORGANISM: Ancylostoma ceylanicum

373 (ix) FEATURE:

374 (A) NAME KEY: Coding Sequence

375 (B) LOCATION: 21...590

376 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

377 GAATTCAC TAATCAACA ATG GCG GTG CTT TAT TCA GTA GCA ATA GCG 50
 378 Met Ala Val Leu Tyr Ser Val Ala Ile Ala
 379 1 5 10
 380 TTA CTA CTG GTA TCA CAA TGC AGT GGG AAA CCG AAC AAT GTG ATG ACT 98
 381 Leu Leu Leu Val Ser Gln Cys Ser Gly Lys Pro Asn Asn Val Met Thr

VERIFICATION SUMMARY

DATE: 06/12/2001

PATENT APPLICATION: US/09/498,556

TIME: 15:06:49

Input Set : N:\Crif3\RULE60\09498556.txt

Output Set: N:\CRF3\06122001\I498556.raw

L:49 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:50 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]

L:2909 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:66
L:2931 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:67
L:2953 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:68
L:3155 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:78
L:3185 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:79
L:3210 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:80
L:3235 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:81
L:3260 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:82
L:3283 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:83
L:3312 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:84
L:3336 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:85
L:3362 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:86
L:3383 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:87
L:3451 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:118
L:3479 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:119
L:4163 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:129
L:4195 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:130
L:4209 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:131
L:4231 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:132
L:4254 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:133
L:4276 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:134
L:4299 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:135
L:4304 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:136
L:4326 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:137
L:4349 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:138
L:4372 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:139
L:4395 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:140
L:4402 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:141
L:4425 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:142
L:4448 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:143
L:4471 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:144
L:4494 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:145
L:4517 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:146
L:4541 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:147
L:4564 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:148
L:4587 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:149
L:4611 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:150
L:4634 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:151
L:4657 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:152
L:4680 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:153
L:4703 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:154
L:4726 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:155
L:4749 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:156

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/498,556

DATE: 06/12/2001

TIME: 15:06:49

Input Set : N:\Crf3\RULE60\09498556.txt

Output Set: N:\CRF3\06122001\I498556.raw

L:4773 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:155
L:4797 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:156
L:4819 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:157
L:4841 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:158